



CLOUD APPLICATION DEVELOPMENT (GROUP 1)

PHASE 1 : ASSIGNMENT NOTEBOOK SUBMISSION

NAME : TAMIL SELVI. P

PROJECT TITLE : Big Data Analysis with IBM Cloud Databases[05]

EMAIL ID : tamiljesus2004@gmail.com

GUIDED BY: MRS.J. HEMALATHA

SPOC NAME : MR. P. VIGNESH

GIT HUB REPOSITORY URL : <https://github.com/TamilSelvip-2004/Bigdata-analysis/settings/keys>

BIG DATA ANALYSIS

INTRODUCTION :

Big data primarily refers to data sets that are too large or complex to be dealt with by *traditional data-processing* application software. Data with many entries offer greater statistical power, while data with higher complexity may lead to a higher false discovery rate . Big data analytics is the process of collecting, examining, and analysing large amounts of data to discover market trends, insights, and patterns that can help companies make better business decisions. Two conspicuous examples are Amazon Prime, which uses *Big Data* analytics to recommend programming for individual users, and Spotify, which does the same to offer personalized music suggestions.

Problem Definition:

The project involves delving into big data analysis using IBM Cloud Databases. The objective is to extract valuable insights from extensive datasets, ranging from climate trends to social patterns. The project includes designing the analysis process, setting up IBM Cloud Databases, performing data analysis, and visualizing the results for business intelligence.

Design Thinking:

1. *Data Selection: Identify the datasets to be analyzed, such as climate data or social media trends.*

- 2. Database Setup: Set up IBM Cloud Databases for storing and managing large datasets.*
- 3. Data Exploration: Develop queries and scripts to explore the datasets, extract relevant information, and identify patterns.*
- 4. Analysis Techniques: Apply appropriate analysis techniques, such as statistical analysis or machine learning, to uncover insights*
- 5. Visualization: Design visualizations to present the analysis results in an understandable and impactful manner.*
- 6. Business Insights: Interpret the analysis findings to derive valuable business intelligence and actionable recommendations.*

What is Big Data?

Big data defined

What exactly is big data?

- The definition of big data is data that contains greater variety, arriving in increasing volumes and with more velocity. This is also known as the three Vs.
- Put simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can't manage them.

Big Data Analytics



purpose of big data :

With big data, you can analyze and assess production, customer feedback and returns, and other factors to reduce outages and anticipate future demands. Big data can also be used to improve decision-making in line with current market demand.

How big data works ?

Big data gives you new insights that open up new opportunities and business models. Getting started involves three key actions:

1. Integrate

Big data brings together data from many disparate sources and applications. Traditional data integration mechanisms, such as extract, transform, and load (ETL) generally aren't up to the task. It requires new strategies and technologies to analyze big data sets at terabyte, or even petabyte, scale.

During integration, you need to bring in the data, process it, and make sure it's formatted and available in a form that your business analysts can get started with.

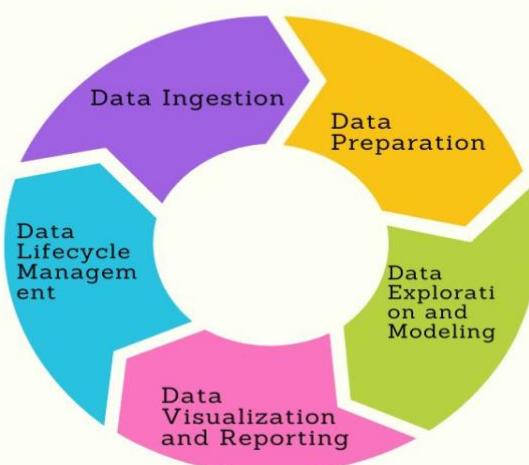
2. Manage

Big data requires storage. Your storage solution can be in the cloud, on premises, or both. You can store your data in any form you want and bring your desired processing requirements and necessary process engines to those data sets on an on-demand basis. Many people choose their storage solution according to where their data is currently residing. The cloud is gradually gaining popularity because it supports your current compute requirements and enables you to spin up resources as needed.

3. Analyze

Your investment in big data pays off when you analyze and act on your data. Get new clarity with a visual analysis of your varied data sets. Explore the data further to make new discoveries. Share your findings with others. Build data models with machine learning and artificial intelligence. Put your data to work.

Life Cycle Phases of Big Data Analytics



Here is a diagram representing this process:

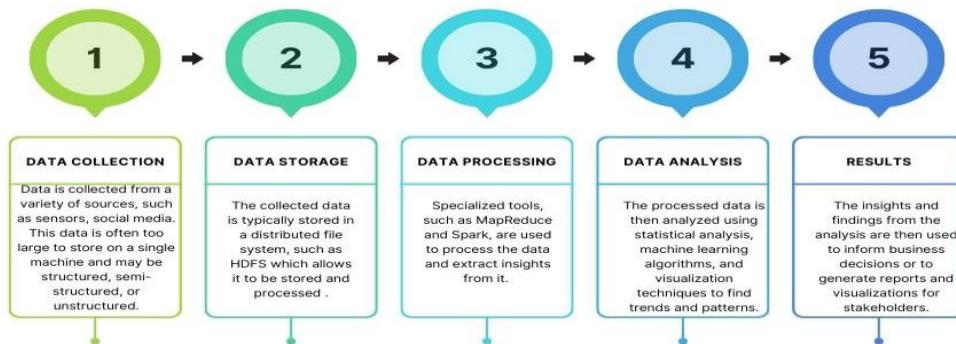
Life Cycle Phases of Big Data Analytics

The following are the phases in the life cycle of big data analytics in brief: 3

- **Data Ingestion:** This is the process of collecting, extracting, and loading data from various sources into a centralized data repository.
- **Data Preparation:** This is the cleaning, transforming, and preparing of data for analysis.
- **Data Exploration and Modeling:** This is the process of using various analytical techniques and tools to uncover patterns and insights in the data.
- **Data Visualization and Reporting:** This is the process of using visual aids to communicate the findings from the data analysis.
- **Data Lifecycle Management:** This is the process of managing the data throughout its lifecycle, from ingestion to visualization and reporting. The data can be collected from various sources like IoT devices, Social Media, Images and

Videos, Applications, and RFID and can be stored and integrated from diverse sources like NoSQL Data Bases, Data Warehouses, Data Lakes.

How does big data analytics works? Management



Applications of Big Data Analytics

There are a number of different applications for big data analytics. Retailers, for example, can use it to track customer behavior and preferences in order to stock their shelves better and design targeted marketing campaigns. Big data analytics can also be used in the healthcare industry to predict disease outbreaks, track the spread of infections, and develop personalized treatments. And in the financial sector, big data analytics is used for fraud detection, risk management, and portfolio optimization. Big data analytics has a wide range of applications across many industries. Here are a few examples:

- Finance: Big data analytics is used in the financial industry to identify trends and patterns in financial data, such as stock prices and market movements. It is also used to detect fraudulent activity and to develop new financial products and services.
- Healthcare: Big data analytics is used in healthcare to improve patient care and outcomes. For example, it can be used to identify trends and patterns in patient data, such as medical history, demographics, and treatment outcomes. This can help healthcare providers to identify risk factors and tailor treatment plans to individual patients.
- Retail: Big data analytics is used in the retail industry to improve customer experience and increase sales. For example, it can be used to analyze customer data, such as purchase history and browsing behavior, to personalize recommendations and targeted marketing campaigns.
- Manufacturing: Big data analytics is used in manufacturing to improve efficiency and reduce costs. For example, it can be used to analyze data from sensors on factory equipment to identify maintenance issues before they become problems.

Future of Big Data Analytics

- The future of big data analytics is looking very bright. With the rapid expansion of data and the need for businesses to make better decisions, the demand for big data analytics is only going to continue to grow.
- A few factors will drive the future of big data analytics. First, there will be an increasing need for real-time insights. This means that businesses will need to be able to analyze data as it's being generated. Second, there will be a growing need for predictive analytics. This means that businesses will need to be able to use historical data to predict future trends. And finally, there will be a need for more advanced analytics, such as machine learning and artificial intelligence.
- With the right tools and technology in place, big data analytics will continue to evolve and become even more powerful.
- The future of big data analytics is looking very bright. As data sets continue to grow larger and more complex, the need for big data analytics will only become more apparent.

- In the future, we can expect to see more businesses using big data analytics to make better decisions, improve their products and services, and save money. We can also expect to see more big data tools and technologies being developed to make the process of big data analytics easier and more efficient.
- Big data analytics is an evolving field, and it's constantly changing and evolving as new technologies are developed. One of the most exciting technologies on the horizon is artificial intelligence. AI can automate the data analysis process, making it even easier to uncover hidden patterns and insights.
- Another promising area of development is edge computing. This involves bringing data processing and storage closer to the edge of the network, where data is collected. This can help to reduce latency and improve performance.
- Finally, quantum computing is also starting to make its way into the world of big data analytics. Quantum computers can process huge amounts of data much faster than traditional computers, which could potentially revolutionize how businesses analyze data.

- So if you're looking for a challenging, interesting, and in-demand career, then a career in big data analytics might be the perfect fit for you.

Conclusion

- The fact is, big data analytics is not just a passing fad. It's a game changer that is only going to become more important in the years to come. Businesses that don't unlock the potential of big data analytics are going to be left behind.
- So what are you waiting for? Start exploring the possibilities of big data analytics today and see how you can improve your business performance. By now, you should know what big data analysts do, the skills they need, and the career opportunities available to them. The world of big data is growing rapidly, and analysts who are able to harness the power of big data analytics will be in high demand. So what are you waiting for? Start learning today and see if a career in big data analytics is right

GitHub Repository URL: <https://github.com/TamilSelvip-2004/Bigdata-analysis/settings/keys>